View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Hermitian Pencils with a Cubic Minimal Polynomial

Helene Shapiro
Department of Mathematics
Swarthmore College
Swarthmore, Pennsylvania 19081

Submitted by Thomas J. Laffey

ABSTRACT

Let A be an $n\times n$ complex matrix and write $A=H+iK$, where H and K are Hermitian matrices. We show that if the minimal polynomial of the pencil $xH+yK$ has degree 3, then there is a unitary matrix U such that $U^{-1}AU$ is block diagonal with blocks of size 3×3 or smaller. This is a special case of a conjecture made by Kippenhahn in 1951.

1. NOTATION AND INTRODUCTION

We work over $\mathbb{C}$, the field of complex numbers, and view $n\times n$ matrices as linear transformations acting on the n dimensional vector space, $\mathbf{V}$, of $n\times 1$ column vectors over $\mathbb{C}$. If $x=(x_1,\dots,x_n)^{\mathsf{T}}$ and $y=(y_1,\dots,y_n)^{\mathsf{T}}$ are two vectors in $\mathbf{V}$, then $\langle x,y\rangle$ denotes the usual inner product: $\langle x,y\rangle=\sum_{i=1}^{n}x_i\bar{y}_i$, where the bar denotes complex conjugation. Let A be an $n\times n$ matrix and write $A=H+iK$, where $H=(A+A^*)/2$ and $K=(A-A^*)/2i$ are Hermitian. We use A^* for the conjugate transpose of A. We use I_k for the $k\times k$ identity matrix, but omit the subscript k if the size of I is clear. We use 0 for the zero matrix.

We say A is *unitarily reducible* if U^*AU is block diagonal for some unitary matrix U. We say a matrix is $D(n_1,n_2,\dots,n_p)$ if it is block diagonal with p diagonal blocks of sizes $n_1\times n_1, n_2\times n_2,\dots,n_p\times n_p$. Recall that $\langle x,Ay\rangle=\langle A^*x,y\rangle$ for all $x,y\in V$. Using this, one can show that A is unitarily reducible to a matrix which is $D(n_1,\dots,n_p)$ if and only if $\mathbf{V}$ can be decomposed as a direct sum, $\mathbf{V}=\mathbf{V}_1\oplus\mathbf{V}_2\oplus\cdots\oplus\mathbf{V}_p$, such that $\mathbf{V}_i$ is a common, invariant subspace of A and A^* of dimension n_i, for each $i=1,\dots,p$ (Specht [12], or see [1] or [10]). Note that $\mathbf{V}_i$ is invariant under both A and A^* if and only if it is invariant under both H and K.

© Elsevier Science Publishing Co., Inc., 1982
52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/070081-23$2.75

The *pencil generated by* H *and* K is $xH + yK = \{rH + sK | r, s \in \mathbb{C}\}$. The *characteristic polynomial of* $xH + yK$ is $\det(\lambda I - xH - yK)$, which we refer to as $f(x, y, \lambda)$. Note that $f(x, y, \lambda)$ is homogeneous of degree n in the variables x, y, and λ. The characteristic polynomial, $\det(\lambda I + xH + yK)$, of $-(xH + yK)$ was studied by Murnaghan [6] and Kippenhahn [3] in connection with the numerical range of $A = H + iK$. See Fiedler [2] for a recent discussion. Motzkin and Taussky [4, 5] studied the characteristic polynomial of a general pencil of matrices in their work on pairs of matrices with property L. Among other things they proved the following result.

THEOREM (Motzkin, Taussky [4]). *Let* H *and* K *be Hermitian matrices. Then* $f(x, y, \lambda) = \det(\lambda I - xH - yK)$ *factors over* $\mathbb{C}$ *into* n *linear factors if and only if* $HK = KH$.

Thus we see A is normal and hence unitarily reducible to a diagonal matrix if and only if $f(x, y, \lambda)$ factors over $\mathbb{C}$ into n linear factors.

If A is unitarily reducible to a matrix which is $D(n_1, \ldots, n_p)$, then $f(x, y, \lambda)$ must factor into p factors of degrees $n_1, \ldots, n_p$. More precisely, if $U^*AU = A_1 \oplus \cdots \oplus A_p$, where A_j is $n_j \times n_j$, and we write $A_j = H_j + iK_j$ where H_j and K_j are Hermitian, then $f(x, y, \lambda) = \prod_{j=1}^{p} \det(\lambda I - xH_j - yK_j)$. However, examples show that $f(x, y, \lambda)$ may factor even when A is not unitarily reducible [3, pp. 205–206; 8, pp. 205–206]. Kippenhahn [3, p. 212]) made the following conjecture.

CONJECTURE (Kippenhahn [3, p. 212]). Let $f(x, y, \lambda) = \det(\lambda I - xH - yK) = [\pi_1(x, y, \lambda)]^{r_1}[\pi_2(x, y, \lambda)]^{r_2} \cdots [\pi_s(x, y, \lambda)]^{r_s}$, where $\pi_1, \ldots, \pi_s$ are distinct irreducible polynomials and the r_i's are positive integers. Then if $r_j > 1$ for some $1 \leqslant j \leqslant s$, the matrix A is unitarily reducible.

The minimal polynomial, denoted $m(x, y, \lambda)$, of the pencil $xH + yK$ is defined in the usual way [3, p. 211; 8, pp. 207–208]. If $f(x, y, \lambda)$ is factored as above, then $m(x, y, \lambda) = \pi_1(x, y, \lambda)\pi_2(x, y, \lambda) \cdots \pi_s(x, y, \lambda)$ [3, p. 211; 8, p. 208]. Thus an alternative way to state the conjecture is that if the degree of the minimal polynomial is less than n, then A is unitarily reducible. Suppose the conjecture is true and $m(x, y, \lambda)$ has degree t. Let $U^*AU = A_1 \oplus \cdots \oplus A_r$, where $A_j = H_j + iK_j$ with H_j and K_j Hermitian. Then $m(x, y, xH_j + yK_j) = 0$, so the minimal polynomial of $xH_j + yK_j$ divides $m(x, y, \lambda)$ and hence has degree at most t. By applying the conjecture to $A_1, A_2, \ldots, A_r$, we can unitarily reduce the blocks until A has been unitarily reduced to a matrix which is $D(n_1, \ldots, n_p)$ with $n_j \leqslant t$ for all $j = 1, \ldots, p$.

Kippenhahn gave a proof for the case where $m(x, y, \lambda)$ has degree two ([3, p. 212], or see [9]). In [8] it was shown that if $f(x, y, \lambda)$ has a linear factor of multiplicity greater than $n/3$, then H and K have a common eigenvector, and hence A is unitarily reducible.

In this paper we verify Kippenhahn's conjecture for the case where the minimal polynomial of $xH + yK$ has degree 3. We show that if $m(x, y, \lambda)$ has degree 3, then A is unitarily reducible to a matrix which is $D(n_1, \dots, n_p)$ with $n_j \leqslant 3$ for $j = 1, \dots, p$.

In Section 2 we show that if $m(x, y, \lambda)$ has degree t, then H and K may be put into a nice block form, consisting of t^2 blocks, via a simultaneous, unitary similarity. In Section 3 we examine the case $t = 3$ more closely to obtain the main result. We present some examples in Section 4. In Section 5 we show Kippenhahn's conjecture holds if $n \leqslant 5$.

2. A BLOCK FORM FOR H AND K

When discussing the pencil $xH + yK$, we may also use the nonhomogeneous form $H + zK$, where z is a complex variable. Note that $\det(\lambda I - (H + zK)) = f(1, z, \lambda)$. The eigenvalues of $H + zK$ are functions of the complex variable z. A theorem of Rellich says these functions are analytic in a neighborhood of 0.

THEOREM (Rellich [7]). *If H and K are Hermitian matrices, then the eigenvalues of $H + zK$ can be expanded in power series about the origin.*

We let $\lambda_1(z), \lambda_2(z), \dots, \lambda_n(z)$ denote the eigenvalues of $H + zK$. Since $H + zK$ may have multiple eigenvalues, the λ_is are not always distinct functions. The eigenvalues of H are $\lambda_1(0), \lambda_2(0), \dots, \lambda_n(0)$; we set $h_i = \lambda_i(0)$. The λ_is are differentiable at 0; we set $k_i = \lambda_i'(0)$. Thus, by Rellich's theorem, in a neighborhood of 0 we have

$$\lambda_i(z) = h_i + k_i z + b_{i2} z^2 + b_{i3} z^3 + \cdots$$

for $i = 1, \dots, n$. Since H and K are Hermitian, $H + zK$ has real eigenvalues whenever z is real. Hence, the coefficients $h_i, k_i, b_{i2}, b_{i3}, \dots$ are all real numbers.

The theorem below, and its proof, are similar to a result of Taussky [13].

THEOREM 1. *Let H and K be $n \times n$ Hermitian matrices, and let $\lambda_1(z), \lambda_2(z), \dots, \lambda_t(z)$ be the distinct eigenvalues of $H + zK$. Let m_i be the multiplicity of $\lambda_i(z)$ for $i = 1, \dots, t$. Let $h_i = \lambda_i(0)$, and assume the numbers*

$h_1,\dots,h_t$ are distinct. Then there is a unitary matrix U such that

$$U^*HU=\begin{pmatrix} h_1I_{m_1} & 0 & \cdots & 0 \\ 0 & h_2I_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_tI_{m_t} \end{pmatrix}$$
$$= h_1I_{m_1}\oplus h_2I_{m_2}\oplus\cdots\oplus h_tI_{m_t}$$

and

$$U^*KU=\begin{pmatrix} k_1I_{m_1} & K_{12} & \cdots & K_{1t} \\ K_{12}^* & k_2I_{m_2} & \cdots & K_{2t} \\ \vdots & \vdots & & \vdots \\ K_{1t}^* & K_{2t}^* & \cdots & k_tI_{m_t} \end{pmatrix}$$

where $k_i=\lambda_i'(0)$ for $i=1,\dots,t$ and the partition of U^*KU into blocks is conformal with that of U^*HU—i.e., the block K_{ij} has size $m_i\times m_j$ for $1\leqslant i<j\leqslant n$.

Proof. Since H is Hermitian, it can be unitarily diagonalized, so we can find a unitary matrix U such that U^*HU is in the desired form. Partition the matrix U^*KU into blocks which are conformal to the block structure of H:

$$U^*KU=\begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1t} \\ K_{12}^* & K_{22} & \cdots & K_{2t} \\ \vdots & \vdots & & \vdots \\ K_{1t}^* & K_{2t}^* & \cdots & K_{tt} \end{pmatrix},$$

where K_{ij} is $m_i\times m_j$ for $1\leqslant i<j\leqslant t$. Without loss of generality, we can assume U^*HU and U^*KU are the original H and K. Henceforth we will call U^*HU and U^*KU by H and K.

We now show $K_{ii}=\lambda_i'(0)I_{m_i}=k_iI_{m_i}$ for each $i=1,\dots,t$.

For convenience of notation we assume $i=1$; the same argument holds for any $i=1,\dots,t$. Since $H+zK$ is Hermitian for every real value of z, and $\lambda_1(z)$ is an eigenvalue of $H+zK$ of multiplicity m_1, the rank of $H+zK-\lambda_1(z)I$ is at most $n-m_1$, whenever z is real. Hence, the determinant of any $(n-m_1+1)\times(n-m_1+1)$ minor of $H+zK-\lambda_1(z)I$ is zero for every real value of z. Since the determinant of a minor of $H+zK-\lambda_1(z)I$ is an analytic function of z in a neighborhood of 0, we see that the determinant of every minor of size $(n-m_1+1)\times(n-m_1+1)$ must be the zero function. We now look at the minor of $H+zK-\lambda_1(z)I$ formed from row i (where $1\leqslant i\leqslant m_1$), column j (where $1\leqslant j\leqslant m_1$), and the last $n-m_1$ rows and columns of $H+zK-\lambda_1(z)I$. We denote the i,j entry of K_{11} by $(K_{11})_{ij}$, the ith row of K_{1r} by $(K_{1r})_i$, and the jth column of K_{1r}^* by $(K_{1r}^*)^j$, for $r=1,\dots,t$.

If $i\neq j$, we have

$$\det\begin{pmatrix} (K_{11})_{ij}z & (K_{12})_i z & \cdots & (K_{1t})_i z \\ (K_{12}^*)^j z & [h_2-\lambda_1(z)]I_{m_2}+K_{22}z & \cdots & K_{2t}z \\ \vdots & \vdots & & \vdots \\ (K_{1t}^*)^j z & K_{2t}^* z & \cdots & [h_t-\lambda_1(z)]I_{m_t}+K_{tt}z \end{pmatrix}$$

$$=0.$$

We differentiate both sides and evaluate at $z=0$. To differentiate the left hand side we must sum all of the determinants in which one row is differentiated. Now, if the first row is not differentiated, then, upon substituting $z=0$, it will become a row of zeros and contribute nothing to the sum. Hence, we need only consider the term in which the first row is differentiated. After substituting $z=0$ we have

$$\det\begin{pmatrix} (K_{11})_{ij} & (K_{12})_i & \cdots & (K_{1t})_i \\ 0 & (h_2-h_1)I_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (h_t-h_1)I_{m_t} \end{pmatrix}=0.$$

Since $h_i-h_1\neq 0$ for $i>1$, we see $(K_{11})_{ij}=0$ whenever $i\neq j$.

If $i=j$, we have

$$\det\begin{pmatrix} h_1+(K_{11})_{ii}z-\lambda_1(z) & (K_{12})_i z & \cdots & (K_{1t})_i z \\ (K_{12}^*)^i z & [h_2-\lambda_1(z)]I_{m_2}+K_{22}z & \cdots & K_{2t}z \\ \vdots & \vdots & & \vdots \\ (K_{1t}^*)^i z & K_{2t}^* z & \cdots & [h_t-\lambda_1(z)]I_{m_t}+K_{tt}z \end{pmatrix}=0.$$

We differentiate and set $z=0$. Since $h_1=\lambda_1(0)$, the same argument used above shows

$$\det\begin{pmatrix} (K_{11})_{ii}-\lambda_1'(0) & (K_{12})_i & \cdots & (K_{1t})_i \\ 0 & (h_2-h_1)I_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (h_t-h_1)I_{m_t} \end{pmatrix}=0.$$

Hence $(K_{11})_{ii}=\lambda_1'(0)$ for $i=1,\dots,m_1$. Thus $K_{11}=\lambda_1'(0)I_{m_1}=k_1I_{m_1}$. ■

The next result gives us information about the off-diagonal blocks K_{ij}.

THEOREM 2. *Let H and K be $n\times n$ Hermitian matrices, and suppose the minimal polynomial of $xH+yK$ has degree t. Assume also that H has t distinct eigenvalues, $h_1,\dots,h_t$, where h_i has multiplicity m_i. Then there is a unitary matrix U such that*

$$U^*HU=\begin{pmatrix} h_1I_{m_1} & 0 & \cdots & 0 \\ 0 & h_2I_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_tI_{m_t} \end{pmatrix}$$

$$=h_1I_{m_1}\oplus h_2I_{m_2}\oplus\cdots\oplus h_tI_{m_t}$$

and

$$U^*KU=\begin{pmatrix} k_1I_{m_1} & K_{12} & \cdots & K_{1t} \\ K_{12}^* & k_2I_{m_2} & \cdots & K_{2t} \\ \vdots & \vdots & & \vdots \\ K_{1t}^* & K_{2t}^* & \cdots & k_tI_{m_t} \end{pmatrix}$$

where K_{ij} is of size $m_i\times m_j$. Furthermore, if we set $a_{ij}=1/(h_i-h_j)$ when $i\neq j$, and $a_{ii}=-\lambda_i''(0)/2$ for $i,j=1,\dots,t$, then the following t equations *hold*:

$$(1)\ a_{11}I_{m_1} + a_{12}K_{12}K_{12}^* + a_{13}K_{13}K_{13}^* + a_{14}K_{14}K_{14}^* + \cdots + a_{1t}K_{1t}K_{1t}^* = 0,$$

$$(2)\ a_{21}K_{12}^*K_{12} + a_{22}I_{m_2} + a_{23}K_{23}K_{23}^* + a_{24}K_{24}K_{24}^* + \cdots + a_{2t}K_{2t}K_{2t}^* = 0,$$

$$(3)\ a_{31}K_{13}^*K_{13} + a_{32}K_{23}^*K_{23} + a_{33}I_{m_3} + a_{34}K_{34}K_{34}^* + \cdots + a_{3t}K_{3t}K_{3t}^* = 0,$$

$$(4)\ a_{41}K_{14}^*K_{14} + a_{42}K_{24}^*K_{24} + a_{43}K_{34}^*K_{34} + a_{44}I_{m_4} + \cdots + a_{4t}K_{4t}K_{4t}^* = 0,$$

$$\vdots$$

$$(t)\ a_{t1}K_{1t}^*K_{1t} + a_{t2}K_{2t}^*K_{2t} + a_{t3}K_{3t}^*K_{3t} + a_{t4}K_{4t}^*K_{4t} + \cdots + a_{tt}I_{m_t} = 0.$$

Proof. Since the minimal polynomial of $xH + yK$ has degree t, and H has t distinct eigenvalues $h_1,\dots,h_t$, we have t distinct eigenvalues, $\lambda_1(z),\dots,\lambda_t(z)$ for $H + zK$, with $\lambda_i(0) = h_i$ for $i = 1,\dots,t$. Since h_i has multiplicity m_i, the eigenvalue $\lambda_i(z)$ must also have multiplicity m_i. So, by Theorem 1, we may assume H and K are in the desired block form, with $k_i = \lambda_i'(0)$ for $i = 1,\dots,t$. It remains to show that the a_{ij}s and K_{ij}s satisfy Equations (1)–(t).

We will show that Equation (1) holds. The same argument serves to derive the other equations. We have $\lambda_1(z) = h_1 + k_1 z + b_{12}z^2 + b_{13}z^3 + \cdots$, where $b_{12} = \lambda_1''(0)/2$. The rank of $H + zK - \lambda_1(z)I$ is at most $n - m_1$, for every value of z, as we saw in the proof of Theorem 1, and

$$H + zK - \lambda_1(z)I$$

$$= \begin{pmatrix} -\left(b_{12}z^2 + b_{13}z^3 + \dots\right)I_{m_1} & K_{12}z & \cdots & K_{1t}z \\ K_{12}^*z & \left[h_2 + k_2 z - \lambda_1(z)\right]I_{m_2} & \cdots & K_{2t}z \\ \vdots & \vdots & & \vdots \\ K_{1t}^*z & K_{2t}^*z & \cdots & \left[h_t + k_t z - \lambda_t(z)\right]I_{m_t} \end{pmatrix}.$$

We remove a factor of z from each of the first m_1 rows and columns of

$H+zK-\lambda_1(z)$ and call the resulting matrix $R(z)$:

$$R(z)=\begin{pmatrix} -(b_{12}+b_{13}z+\dots)I_{m_1} & K_{12} & \cdots & K_{1t} \\ K_{12}^* & [h_2+k_2z-\lambda_1(z)]I_{m_2} & \cdots & K_{2t}z \\ \vdots & \vdots & & \vdots \\ K_{1t}^* & K_{2t}^*z & \cdots & [h_t+k_tz-\lambda_1(z)]I_{m_t} \end{pmatrix}.$$

If $z \neq 0$, then $R(z)$ and $H+zK-\lambda_1(z)I$ have the same rank. Hence, the determinant of every minor of $R(z)$ of size $(n-m_1+1)\times(n-m_1+1)$ is zero for $z \neq 0$. But the determinant of a minor of $R(z)$ is an analytic function of z. Hence, every minor of $R(z)$ of size $(n-m_1+1)\times(n-m_1+1)$ vanishes identically in z and the rank of $R=R(0)$ is at most $n-m_1$. We have

$$R=R(0)=\begin{pmatrix} -b_{12}I_{m_1} & K_{12} & K_{13} & \cdots & K_{1t} \\ K_{12}^* & (h_2-h_1)I_{m_2} & 0 & \cdots & 0 \\ K_{13}^* & 0 & (h_3-h_1)I_{m_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ K_{1t}^* & 0 & 0 & \cdots & (h_t-h_1)I_{m_t} \end{pmatrix}.$$

Since $h_i - h_1 \neq 0$ for $i=2,\dots,t$, the matrix R has rank $n-m_1$. Thus the null space of R has dimension m_1.

Suppose X is in the null space of R. We write

$$X=\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_t \end{pmatrix},$$

where each X_i is a column vector of length m_i. We then have

$$RX = \begin{pmatrix} -b_{12}X_1 + K_{12}X_2 + K_{13}X_3 + \cdots + K_{1t}X_t \\ K_{12}^*X_1 + (h_2 - h_1)X_2 \\ K_{13}^*X_1 + (h_3 - h_1)X_3 \\ \vdots \\ K_{1t}^*X_1 + (h_t - h_1)X_t \end{pmatrix} = 0.$$

Hence

$$X_j = \frac{1}{h_1 - h_j} K_{1j}^* X_1 \qquad \text{for each} \quad j = 2, \ldots, t.$$

Thus, the vector X is completely determined by the column vector X_1 of length m_1. Since the null space of R has dimension m_1, for any choice of the vector X_1,

$$X = \begin{pmatrix} X_1 \\ \frac{1}{h_1 - h_2} K_{12}^* X_1 \\ \frac{1}{h_1 - h_3} K_{13}^* X_1 \\ \vdots \\ \frac{1}{h_1 - h_t} K_{1t}^* X_1 \end{pmatrix}$$

is in the nulll space of R. Computing RX, we see that for any vector X_1 we have

$$\left(-b_{12} I_{m_1} + \frac{1}{h_1 - h_2} K_{12} K_{12}^* + \frac{1}{h_1 - h_3} K_{13} K_{13}^* + \cdots + \frac{1}{h_1 - h_t} K_{1t} K_{1t}^* \right) X_1 = 0.$$

Hence, if we put $a_{11} = -b_{12} = -\lambda_1''(0)/2$ and $a_{1j} = 1/(h_1 - h_j)$ for $j =$

$2,\ldots,t$, then we obtain

$$a_{11}I_{m_1} + a_{12}K_{12}K_{12}^* + a_{13}K_{13}K_{13}^* + \cdots + a_{1t}K_{1t}K_{1t}^* = 0,$$

which is (1). The other equations are obtained in a similar fashion. ∎

3. $m(x, y, \lambda)$ HAS DEGREE 3

We now assume the minimal polynomial, $m(x, y, \lambda)$, of $xH + yK$ has degree three and prove our main result. We will need some facts about matrix products of the form SS^* and S^*S, which we collect in the lemma below.

LEMMA 1. *Let S be a complex $p \times q$ matrix. Let $\mathbf{W}$ be the vector space of column vectors of length q, and let $\mathbf{V}$ be the vector space of column vectors of length p. We view S as a linear transformation from $\mathbf{W}$ into $\mathbf{V}$ and S^* as a linear transformation of $\mathbf{V}$ into $\mathbf{W}$. Then the following hold:*

(1) *SS^* is a $p \times p$, positive semidefinite Hermitian matrix. S^*S is a $q \times q$, positive semidefinite Hermitian matrix.*

(2) *$SS^*v = 0$ if and only if $S^*v = 0$. $S^*Sw = 0$ if and only if $Sw = 0$.*

(3) *S, S^*, SS^*, and S^*S all have the same rank.*

(4) *If σ is a nonzero eigenvalue of SS^* with multiplicity m and associated eigenspace $\mathbf{V}_\sigma$, then σ is an eigenvalue of S^*S with multiplicity m and eigenspace $S^*(\mathbf{V}_\sigma)$.*

These facts are well known; we omit the proof.

THEOREM 3. *Let H and K be $n \times n$ Hermitian matrices, and let $m(x, y, \lambda)$ be the minimal polynomial of $xH + yK$. Suppose $m(x, y, \lambda)$ has degree three. Then $A = H + iK$ is unitarily reducible to a matrix which is $D(n_1, n_2, \ldots, n_p)$ where $n_j \leqslant 3$ for each $j = 1, \ldots, p$.*

Proof. If we can show that H and K have a nontrivial, common invariant subspace of dimension at most three, then we can complete the argument by using induction on n. For, if H and K have a nontrivial, common invariant subspace of dimension $n_1 \leqslant 3$, then for some unitary matrix U we have $U^*AU = A_1 \oplus A_2$, where A_1 is $n_1 \times n_1$. Writing $A_2 = H_2 + iK_2$, where H_2 and K_2 are Hermitian, we see that the minimal polynomial of $xH_2 + yK_2$ divides $m(x, y, \lambda)$ because $m(x, y, xH_2 + yK_2) = 0$. Hence, the minimal polynomial of $xH_2 + yK_2$ has degree less than or equal to 3. Since A_2 is $(n - n_1) \times (n -$

n_1), by induction we can assume A_2 is unitarily reducible to a matrix which is $D(n_2,\dots,n_p)$ with $n_j \leqslant 3$ for $j=2,\dots,p$. Hence A is unitarily reducible to a matrix which is $D(n_1,\dots,n_p)$ with $n_j \leqslant 3$ for $j=1,\dots,p$. Thus it suffices to show H and K have a nontrivial, common invariant subspace of dimension at most three.

We first show that, without loss of generality, we may assume H has three distinct eigenvalues. Since $m(x,y,\lambda)$ has degree three, $H+zK$ has three distinct eigenvalues, $\lambda_1(z)$, $\lambda_2(z)$, and $\lambda_3(z)$. For some real number z_0, the numbers $\lambda_1(z_0)$, $\lambda_2(z_0)$, and $\lambda_3(z_0)$ are distinct and $H+z_0K$ has three distinct eigenvalues. Notice that H and K generate the same pencil as $H+z_0K$ and K do, and replacing H by $H+z_0K$ in $f(x,y,\lambda)=\det(\lambda I - xH - yK)$ has the same effect on $f(x,y,\lambda)$ and $m(x,y,\lambda)$ as a linear change of variable. So, if we replace H by $H+z_0K$, we do not change the pencil $xH+yK$ or the form of $f(x,y,\lambda)$ and $m(x,y,\lambda)$. Also, a subspace of $\mathbf{V}$ is invariant under both H and K if and only if it is invariant under both $H+z_0K$ and K, so U^*AU is $D(n_1,\dots,n_p)$ if and only if $U^*[(H+z_0K)+iK]U$ is $D(n_1,\dots,n_p)$. So, without loss of generality, we may assume the original matrix, H, has three distinct eigenvalues, h_1, h_2, and h_3, of multiplicities m_1, m_2, and m_3, respectively. We order the h_is so that $m_1 \geqslant m_2 \geqslant m_3$. (This will be used in one case of the argument.) As before, we have $h_i = \lambda_i(0)$ and $k_i = \lambda_i'(0)$ for $i=1,2,3$.

By Theorem 2, we may assume

$$H=\begin{pmatrix} h_1I_{m_1} & 0 & 0\\ 0 & h_2I_{m_2} & 0\\ 0 & 0 & h_3I_{m_3}\end{pmatrix} \quad\text{and}\quad K=\begin{pmatrix} k_1I_{m_1} & B & C\\ B^* & k_2I_{m_2} & D\\ C^* & D^* & k_3I_{m_3}\end{pmatrix}$$

where B is $m_1\times m_2$, while C is $m_1\times m_3$ and D is $m_2\times m_3$. Furthermore, for some constants α_i, β_i, and γ_i, with $i=1,2,3$ and $\alpha_i\beta_i \neq 0$, the following three equations hold:

$$\alpha_1BB^* + \beta_1CC^* = \gamma_1 I_{m_1}, \tag{1}$$

$$\alpha_2B^*B + \beta_2DD^* = \gamma_2 I_{m_2}, \tag{2}$$

$$\alpha_3C^*C + \beta_3D^*D = \gamma_3 I_{m_3}. \tag{3}$$

We now show that H and K have a nontrivial, common invariant subspace of dimension at most three. The block form of H and K suggests that we try to find column vectors X of length m_1, Y of length m_2, and Z of length m_3

such that the subspace spanned by

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ Y \\ 0_{m_3} \end{pmatrix}, \text{ and } \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ Z \end{pmatrix}$$

is invariant under both H and K. Here, 0_{m_i} means a column of m_i zeros, for $i = 1,2,3$. This space is clearly invariant under H for any choice of X, Y, and Z. We study the matrices B, C, and D to see how X, Y, and Z may be chosen so that this space is also invariant under K.

In addition to Equations (1)–(3), we need

$$\alpha_4 BB^* + \beta_4 CC^* + BDC^* + CD^*B^* = \gamma_4 I_{m_1} \tag{4}$$

for some constants $\alpha_4, \beta_4, \gamma_4$. This is derived as follows. Let

$$m(x,y,\lambda) = \lambda^3 + (ax+by)\lambda^2 + q(x,y)\lambda + c(x,y),$$

where a and b are real numbers, $q(x,y)$ is a homogeneous polynomial of degree two, and $c(x,y)$ is a homogeneous polynomial of degree 3. Since $m(0,1,K) = 0$ we have

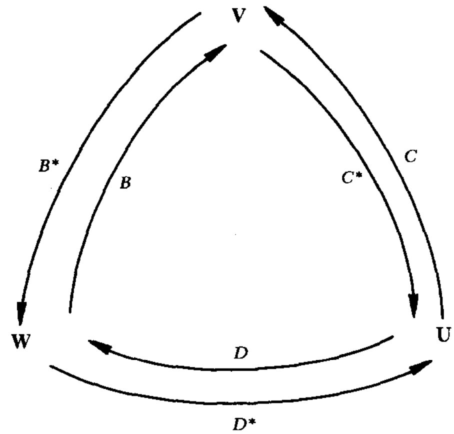
$$K^3 + bK^2 + q(0,1)K + c(0,1)I = 0. \tag{5}$$

Using block multiplication to compute K^2 and K^3, we find the first $m_1 \times m_1$ diagonal block of the matrix expression on the left-hand side of (5). This yields

$$(2k_1 + k_2 + b)BB^* + (2k_1 + k_3 + b)CC^* + BDC^* + CD^*B^* + \left[k_1^3 + bk_1^2 + q(0,1)k_1 + c(0,1)\right]I_{m_1} = 0. \tag{6}$$

Hence, there exist constants α_4, β_4, and γ_4 such that (4) holds.

Now let $\mathbf{V}$ be the vector space of column vectors of length m_1, while $\mathbf{W}$ is the space of column vectors of length m_2, and $\mathbf{U}$ the space of column vectors of length m_3. Figure 1 is helpful in keeping track of $\mathbf{V}$, $\mathbf{W}$, and $\mathbf{U}$. Let $b_1,\dots,b_r$ be the distinct eigenvalues of BB^*, and let $\mathbf{V}_i$ be the eigenspace of BB^* associated with b_i for $i = 1,\dots,r$. Then we have $\mathbf{V} = \mathbf{V}_1 \oplus \mathbf{V}_2 \oplus \cdots \oplus \mathbf{V}_r$.



$\dim \mathbf{V} = m_1,$	B is $m_1 \times m_2,$	B^* is $m_2 \times m_1,$
$\dim \mathbf{W} = m_2,$	C is $m_1 \times m_3,$	C^* is $m_3 \times m_1,$
$\dim \mathbf{U} = m_3,$	D is $m_2 \times m_3,$	D^* is $m_3 \times m_2.$

FIG. 1

Recall Equations (1)–(3):

$$\alpha_1 BB^* + \beta_1 CC^* = \gamma_1 I_{m_1}, \tag{1}$$

$$\alpha_2 B^*B + \beta_2 DD^* = \gamma_2 I_{m_2}, \tag{2}$$

$$\alpha_3 C^*C + \beta_3 D^*D = \gamma_3 I_{m_3}. \tag{3}$$

Recall also that $\alpha_j\beta_j \neq 0$ for $j = 1,2,3$. Set $c_i = (\gamma_1 - \alpha_1 b_i)/\beta_1$ and $d_i = (\gamma_3 - \alpha_3 c_i)/\beta_3$. From Equation (1) we see that V_i is the eigenspace of CC^* associated with the eigenvalue c_i. Now, let

$$\mathbf{U}_i = \{u \in \mathbf{U} | C^*Cu = c_i u\}.$$

By (3),

$$\mathbf{U}_i = \{u \in \mathbf{U} | D^*Du = d_i u\}.$$

Let

$$\mathbf{W}_i = \{w \in \mathbf{W} | DD^*w = d_i w\}.$$

By (2),

$$\mathbf{W}_i = \{w \in \mathbf{W} | B^*Bw = bw\}, \qquad \text{where} \quad b = \frac{\gamma_2 - \beta_2 d_i}{\alpha_2}.$$

Let $\mathbf{V}_b = \{v \in \mathbf{V} | BB^*v = bv\}$. One can easily see that $C^*(\mathbf{V}_i) \subseteq \mathbf{U}_i$, while $D(\mathbf{U}_i) \subseteq \mathbf{W}_i$ and $B(\mathbf{W}_i) \subseteq \mathbf{V}_b$. Hence $BDC^*(\mathbf{V}_i) \subseteq \mathbf{V}_b$. If $X \in \mathbf{V}_i$ and $X \neq 0$, then $BDC^*X = 0$ if and only if $bc_i d_i = 0$, by Lemma 1, parts (2) and (4). If $bc_i d_i \neq 0$, then by part (4) of Lemma 1 we have $C^*(\mathbf{V}_i) = \mathbf{U}_i$, $D(\mathbf{U}_i) = \mathbf{W}_i$, and $B(\mathbf{W}_i) = \mathbf{V}_b$. In this case, $\mathbf{V}_i$, $\mathbf{U}_i$, $\mathbf{W}_i$, and $\mathbf{V}_b$ all have the same dimension and b must equal b_j for some eigenvalue b_j of BB^*. We have $b = b_i$ if and only if $\alpha_2 b_i + \beta_2 d_i = \gamma_2$. Thus, there are three possibilities:

(7.1) $BDC^*X = 0$ for some nonzero X in $\mathbf{V}_i$ if and only if $bc_i d_i = 0$. In this case, $BDC^*(\mathbf{V}_i) = \{0\}$.

(7.2) If $bc_i d_i \neq 0$ and $\alpha_2 b_i + \beta_2 d_i \neq \gamma_2$ then $BDC^*(\mathbf{V}_i) = \mathbf{V}_j$ for some $j \neq i$.

(7.3) If $bc_i d_i \neq 0$ and $\alpha_2 b_i + \beta_2 d_i = \gamma_2$ then $b = b_i$ and $BDC^*(\mathbf{V}_i) = \mathbf{V}_i$.

Similarly, if $d = (\gamma_2 - \alpha_2 b_i)/\beta_2$ and $c = (\gamma_3 - \beta_3 d)/\alpha_3$, then we again have three possible cases:

(8.1) $CD^*B^*X = 0$ for some nonzero X in V_i if and only if $b_i cd = 0$. In this case, $CD^*B^*(\mathbf{V}_i) = \{0\}$.

(8.2) If $b_i cd \neq 0$ and $\alpha_2 b_i + \beta_2 d_i \neq \gamma_2$ then $CD^*B^*(\mathbf{V}_i) = \mathbf{V}_k$ for some $k \neq i$.

(8.3) If $b_i cd \neq 0$ and $\alpha_2 b_i + \beta_2 d_i = \gamma_2$ then $d = d_i$ and $c = c_i$ and $CD^*B^*(\mathbf{V}_i) = \mathbf{V}_i$.

We now use Equation (4):

$$\alpha_4 BB^* + \beta_4 CC^* + BDC^* + CD^*B^* = \gamma_4 I_{m_1}. \tag{4}$$

For any $X \in \mathbf{V}_i$, we have

$$BDC^*X + CD^*B^*X = (\gamma_4 - \alpha_4 b_i - \beta_4 c_i)X. \tag{9}$$

First, suppose $X \neq 0$ and $BDC^*X = 0$. Then $CD^*B^*X \in \mathbf{V}_i$. If $CD^*B^*X \neq 0$, then by (8.1), (8.2), and (8.3) we have $b_i cd \neq 0$ and $\alpha_2 b_i + \beta_2 d_i = \gamma_2$. But then $c = c_i$ and $d = d_i$ so $b_i c_i d_i \neq 0$. This contradicts (7.1), because $b = b_i$ and we assumed $BDC^*X = 0$. So, if $BDC^*X = 0$ for any nonzero X in $\mathbf{V}_i$, then we must have $BDC^*(\mathbf{V}_i) = CD^*B^*(\mathbf{V}_i) = \{0\}$. Similarly, if $CD^*B^*X = 0$, then we must have $CD^*B^*(\mathbf{V}_i) = BDC^*(\mathbf{V}_i) = \{0\}$. Thus, there are three cases to consider:

Case 1. $BDC^*X \neq 0$ for some X in $\mathbf{V}_i$ and $\alpha_2 b_i + \beta_2 d_i = \gamma_2$. Then by (7.1), (7.3), (8.1), and (8.3) we have $b_i c_i d_i \neq 0$ and $BDC^*(\mathbf{V}_i) = CD^*B^*(\mathbf{V}_i) = \mathbf{V}_i$.

Case 2. $BDC^*(\mathbf{V}_i) = CD^*B^*(\mathbf{V}_i) = 0$ for all $i = 1,\dots,r$.

Case 3. If neither case 1 nor case 2 holds then for some i we have $BDC^*X \neq 0$ for some X in $\mathbf{V}_i$ and $\alpha_2 b_i + \beta_2 d_i \neq \gamma_2$. Then by (7.1) and (7.2) we have $BDC^*(\mathbf{V}_i) = \mathbf{V}_j$ for some $i \neq j$. From (9), and the fact that $\mathbf{V} = \mathbf{V}_1 \oplus \mathbf{V}_2 \oplus \cdots \oplus \mathbf{V}_r$ is a direct sum, we must have $CD^*B^*X = -BDC^*X$ for all X in $\mathbf{V}_i$ and $\gamma_4 - \alpha_4 b_i - \beta_4 c_i = 0$.

We will show that in each of the first two cases, H and K have a nontrivial, common, invariant subspace of dimension at most three. We then complete the proof of Theorem 3 by showing that case 3 cannot arise.

In case 1, let $X \in \mathbf{V}_i$ be an eigenvector of BDC^* with nonzero eigenvalue λ. Thus, $BDC^*X = \lambda X$. Multiply both sides on the left by B^* to get $B^*B(DC^*X) = \lambda B^*X$. Since $DC^*X \in \mathbf{W}_i$, we have $B^*B(DC^*X) = b_i DC^*X$. Hence we have

$$DC^*X = \frac{\lambda}{b_i} B^*X. \tag{10}$$

Now multiply both sides of (10) on the left by D^*. Since $C^*X \in \mathbf{U}_i$, we have $D^*D(C^*X) = d_i C^*X = (\lambda / b_i) D^*B^*X$. Thus, we obtain

$$D^*B^*X = \frac{b_i d_i}{\lambda} C^*X. \tag{11}$$

Using Equations (10) and (11) plus the fact that $BB^*X = b_i X$ and $CC^*X = c_i X$, one can see that the space spanned by the three vectors

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ B^*X \\ 0_{m_3} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ C^*X \end{pmatrix}$$

is invariant under K.

In case 2, we have $BDC^*X = CD^*B^*X$ for every X in $\mathbf{V}_i$ for each $i = 1,\dots,r$. Hence, $BDC^* = CD^*B^* = 0$. If $b_i = 0$ for some i, let X be a nonzero vector of $\mathbf{V}_i$. Then by part (2) of Lemma 1 we have $B^*X = 0$. Also, $C^*X \in \mathbf{U}_i$, so $D^*D(C^*X) = d_i C^*X$. Hence the space spanned by the three vectors

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ C^*X \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0_{m_1} \\ DC^*X \\ 0_{m_3} \end{pmatrix}$$

is invariant under K. Similarly, if $c_i = 0$ for some i, then, if X is any nonzero vector of V_i, the space spanned by

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ B^*X \\ 0_{m_3} \end{pmatrix}, \text{ and } \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ D^*B^*X \end{pmatrix}$$

is invariant under K. If $b_i c_i \neq 0$ for any choice of $i = 1, \dots, r$, then the $m_1 \times m_1$ matrices BB^* and CC^* are nonsingular. By Lemma 1, part (3), both B and C have rank m_1. Since B is $m_1 \times m_2$ and C is $m_1 \times m_3$, this implies $m_1 \leqslant m_2$ and $m_1 \leqslant m_3$. However, we assumed at the start that $m_1 \geqslant m_2 \geqslant m_3$. This forces $m_1 = m_2 = m_3$. Hence, B and C are square, nonsingular matrices and $BDC^* = 0$ implies $D = 0$. Hence, for any nonzero X in $\mathbf{V}_i$, the space spanned by the vectors

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ B^*X \\ 0_{m_3} \end{pmatrix}, \text{ and } \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ C^*X \end{pmatrix}$$

is invariant under K.

It remains to deal with case 3. We shall assume that case 3 does occur and show that this leads to a contradiction.

In case 3, we have $CD^*B^*X = -BDC^*X$ for all X in $\mathbf{V}_i$. When $X \neq 0$, the vector CD^*B^*X is nonzero and lies in $\mathbf{V}_j$, for some $j \neq i$. Let X be any nonzero vector of $\mathbf{V}_i$. Then DC^*X is in $\mathbf{W}_i$, so we have

$$B^*B(DC^*X) = b_j DC^*X. \tag{12}$$

Since $CD^*B^*X = -BDC^*X$, we have $C^*BDC^*X = -C^*C(D^*B^*X)$. By Lemma 1 (4) and Equations (1)–(3), we know D^*B^*X is an eigenvector of C^*C, so

$$C^*BDC^*X = \lambda D^*B^*X \qquad \text{for some scalar } \lambda. \tag{13}$$

Since $C^*X \in U_i$, we know C^*X is an eigenvector of D^*D; by (2) and Lemma 1, we know B^*X is an eigenvector of DD^*. Using Equations (12) and (13), the fact that $CD^*B^*X = -BDC^*X$ for all X in $\mathbf{V}_i$, and the facts of the previous sentence one can verify that the space spanned by the six vectors

$$\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} BDC^*X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ B^*X \\ 0_{m_3} \end{pmatrix},$$

$$\begin{pmatrix} 0_{m_1} \\ DC^*X \\ 0_{m_3} \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ C^*X \end{pmatrix}, \quad \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ D^*B^*X \end{pmatrix}$$

is invariant under H and K. Let $\mathcal{V}$ denote the subspace spanned by these six vectors. First note that all six of these vectors are nonzero; we now show that they are linearly independent. Since $X \in \mathbf{V}_i$ and $BDC^*X \in \mathbf{V}_j$, with $i \neq j$, we know X and BDC^*X are linearly independent. Also, if $DC^*X = \lambda B^*X$ for some scalar λ, then $BDC^*X = \lambda BB^*X \in \mathbf{V}_i$, which contradicts the conditions of case 3. Hence DC^*X and B^*X are linearly independent. Similarly, if $D^*B^*X = \lambda C^*X$, then we have $CD^*B^*X = \lambda CC^*X \in \mathbf{V}_i$, again contradicting the conditions of case 2. Hence the six vectors are linearly independent and $\mathcal{V}$ has dimension six.

For the remainder of the argument we need only consider the action of H and K on the subspace $\mathcal{V}$. Thus, it will suffice to deal with the case where H and K are 6×6. We first normalize the vectors X, B^*X, and C^*X and use the Gram-Schmidt process to construct vectors Y, Z, and T so that the following six vectors form an orthonormal basis for $\mathcal{V}$:

$$v_1 = \frac{1}{\|X\|}\begin{pmatrix} X \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \qquad v_2 = \begin{pmatrix} Y \\ 0_{m_2} \\ 0_{m_3} \end{pmatrix}, \qquad v_3 = \frac{1}{\|B^*X\|}\begin{pmatrix} 0_{m_1} \\ B^*X \\ 0_{m_3} \end{pmatrix},$$

$$v_4 = \begin{pmatrix} 0_{m_1} \\ Z \\ 0_{m_3} \end{pmatrix}, \qquad v_5 = \frac{1}{\|C^*X\|}\begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ C^*X \end{pmatrix}, \qquad v_6 = \begin{pmatrix} 0_{m_1} \\ 0_{m_2} \\ T \end{pmatrix}.$$

Now let $\tilde{H}$ and $\tilde{K}$ be the 6×6 matrices which represent the actions of H and K, respectively, on the subspace $\mathcal{V}$, with respect to the new basis $\{v_1, v_2, v_3, v_4, v_5, v_6\}$. Then we have $\tilde{H} = h_1 I_2 \oplus h_2 I_2 \oplus h_3 I_2$, where I_2 is the 2×2 identity matrix. Computing the products Kv_i, where $i = 1,\dots,6$, we find that $\tilde{K}$ has the following form:

$$\tilde{K}=\left(\begin{array}{cc|cc|cc} k_1 & 0 & k_{13} & k_{14} & k_{15} & k_{16} \\ 0 & k_1 & 0 & k_{24} & 0 & k_{26} \\ \hline k_{31} & k_{32} & k_2 & 0 & k_{35} & k_{36} \\ 0 & k_{42} & 0 & k_2 & k_{45} & k_{46} \\ \hline k_{51} & k_{52} & k_{53} & k_{54} & k_3 & 0 \\ 0 & k_{62} & k_{63} & k_{64} & 0 & k_3 \end{array}\right).$$

Here, the k_{ij}s are undetermined coefficients, while k_1, k_2, and k_3 are from the original matrix K. However, since $\{v_1, v_2, v_3, v_4, v_5, v_6\}$ is an orthonormal basis, $\tilde{H}$ and $\tilde{K}$ are obtained by applying a simultaneous, unitary similarity to H and K, and thus still satisfy the hypothesis of Theorem 3. Since $\tilde{K}$ must be Hermitian, we have $k_{14}=k_{16}=k_{32}=k_{52}=0$. Setting $b_1=k_{13}$, $b_2=k_{24}$, $c_1=k_{15}$, and $c_2=k_{26}$ we then have

$$\tilde{K}=\left(\begin{array}{cc|cc|cc} k_1 & 0 & b_1 & 0 & c_1 & 0 \\ 0 & k_1 & 0 & b_2 & 0 & c_2 \\ \hline \bar{b}_1 & 0 & k_2 & 0 & k_{35} & k_{36} \\ 0 & \bar{b}_2 & 0 & k_2 & k_{45} & k_{46} \\ \hline \bar{c}_1 & 0 & \bar{k}_{35} & \bar{k}_{45} & k_3 & 0 \\ 0 & \bar{c}_2 & \bar{k}_{36} & \bar{k}_{46} & 0 & k_3 \end{array}\right).$$

Furthermore, since $\tilde{H}$ still has the diagonal form specified in Theorems 1 and 2, the off diagonal blocks of $\tilde{K}$ satisfy the equations of Theorem 2. Thus, the three matrices

$$B=\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}, \qquad C=\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}, \quad \text{and} \quad D=\begin{pmatrix} k_{35} & k_{36} \\ k_{45} & k_{46} \end{pmatrix}$$

satisfy equations (1)–(4).

Note that $BB^* = B^*B$ has eigenvalues $|b_1|^2$ and $|b_2|^2$. Since we are in case 3, these eigenvalues must be distinct, i.e., $|b_1| \neq |b_2|$. Also, the eigenspace $\mathbf{V}_1$ is now spanned by $X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, while the eigenspace $\mathbf{V}_2$ is spanned by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We compute[1]

$$BDC^*X = \begin{pmatrix} b_1\bar{c}_1k_{35} \\ b_2\bar{c}_1k_{45} \end{pmatrix} \quad \text{and} \quad CD^*B^*X = \begin{pmatrix} c_1\bar{b}_1\bar{k}_{35} \\ c_2\bar{b}_1\bar{k}_{36} \end{pmatrix}.$$

Now, under the conditions of case 3, $BDC^*X \neq 0$ and $CD^*B^*X \neq 0$, so $b_1 \neq 0$ and $c_1 \neq 0$. Also, $BDC^*X \in \mathbf{V}_2$, so $b_1\bar{c}_1k_{35} = 0$. Hence $k_{35} = 0$. Then from

$$BDC^*X = \begin{pmatrix} 0 \\ b_2\bar{c}_1k_{45} \end{pmatrix} \neq 0$$

we see that $b_2 \neq 0$, and from

$$CD^*B^*X = \begin{pmatrix} 0 \\ c_2\bar{b}_1\bar{k}_{36} \end{pmatrix} \neq 0$$

we see that $c_2 \neq 0$ and $k_{36} \neq 0$. We now compute

$$BDC^*\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1\bar{c}_2k_{36} \\ b_2\bar{c}_2k_{46} \end{pmatrix}.$$

Since $b_1\bar{c}_2k_{36} \neq 0$, this vector is not in $\mathbf{V}_2$ and hence $BDC^*(\mathbf{V}_2) \subseteq \mathbf{V}_1$. Hence $b_2\bar{c}_2k_{46} = 0$. Since $b_2\bar{c}_2 \neq 0$, we have $k_{46} = 0$. Hence

$$D = \begin{pmatrix} 0 & k_{36} \\ k_{45} & 0 \end{pmatrix}.$$

For convenience, we set $r = k_{36}$ and $s = k_{46}$, or

$$D = \begin{pmatrix} 0 & r \\ s & 0 \end{pmatrix}.$$

[1] The argument is the same for $X = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

We now use Equations (1)–(3). Recalling what the coefficients α_i and β_i are from Theorem 2, we have

$$\frac{1}{h_1-h_2}BB^*+\frac{1}{h_1-h_3}CC^*=\gamma_1 I_2, \tag{1}$$

$$\frac{1}{h_2-h_1}B^*B+\frac{1}{h_2-h_3}DD^*=\gamma_2 I_2, \tag{2}$$

$$\frac{1}{h_3-h_1}C^*C+\frac{1}{h_3-h_2}D^*D=\gamma_3 I_2. \tag{3}$$

Adding Equations (1)–(3) together, and using the facts that $BB^*=B^*B$ and $CC^*=C^*C$, we obtain

$$\frac{1}{h_2-h_3}(DD^*-D^*D)=(\gamma_1+\gamma_2+\gamma_3)I_2.$$

Hence DD^*-D^*D is a scalar matrix. However,

$$DD^*-D^*D=\begin{pmatrix}|r|^2-|s|^2 & 0\\ 0 & |s|^2-|r|^2\end{pmatrix}.$$

Hence we must have $|r|=|s|$ and $DD^*=D^*D=|r|^2I_2$. Equation (2) now implies that

$$B^*B=\begin{pmatrix}|b_1|^2 & 0\\ 0 & |b_2|^2\end{pmatrix}$$

must also be a scalar matrix, contradicting the condition that in case 3 the matrix B^*B must have at least two distinct eigenvalues. Hence case 3 cannot occur. This completes the proof of Theorem 3. ■

4. EXAMPLES

It was shown in [9] that if $f(x,y,\lambda)=[g(x,y,\lambda)]^{n/2}$, where $g(x,y,\lambda)$ is irreducible of degree two, then A is unitarily similar to

$$\underbrace{A_1\oplus A_1\oplus\cdots\oplus A_1}_{n/2\,\text{times}},$$

where the 2×2 matrix A_1 is uniquely determined, up to a unitary similarity, by the polynomial $g(x,y,\lambda)$. In fact, a 2×2 matrix $A_1=H_1+iK_1$ is always uniquely determined, up to a unitary similarity, by the characteristic polynomial of xH_1+yK_1 [8, pp. 206–207 or 9, pp. 104–105]. This is not true for 3×3 matrices, as Example 1 of [8] shows.

EXAMPLE 1 [8, pp. 205–206]. Let

$$A=\begin{pmatrix} 0 & i & i \\ i & a+i & i \\ i & i & -a+i \end{pmatrix}=H_1+iK_1$$

and

$$B=\begin{pmatrix} 0 & 0 & 0 \\ 0 & a+i & i\sqrt{3} \\ 0 & i\sqrt{3} & -a+i \end{pmatrix}=H_2+iK_2,$$

where a is any nonzero real number. Then $\det(\lambda I-xH_1-yK_1)=\det(\lambda I-xH_2-yK_2)=\lambda[(\lambda-ax-y)(\lambda+ax-y)-3y^2]$. However, A is not unitarily reducible, while B is $D(1,2)$. Hence, for no unitary matrix U do we have $U^*AU=B$.

Furthermore, even when $f(x,y,\lambda)$ is an irreducible cubic, it does not uniquely determine A up to unitary similarity.

EXAMPLE 2 [11, pp. 28–31]. Let

$$A=\begin{pmatrix} 0 & 1/\sqrt{2} & 2 \\ 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}=H_1+iK_1$$

and

$$B=\begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}=H_2+iK_2.$$

Then

$$\begin{aligned} f(x,y,\lambda)&=\det(\lambda I-xH_1-yK_1)=\det(\lambda I-xH_2-yK_2)\\ &=\lambda^3-\tfrac{5}{4}(x^2+y^2)\lambda-\tfrac{1}{4}x(x^2+y^2). \end{aligned}$$

If $f(x, y, \lambda)$ has a linear factor, it must be of the form $\lambda - \alpha x - \beta y$, where $\alpha + i\beta$ is an eigenvalue of A, while α is an eigenvalue of H and β is an eigenvalue of K. Since 0 is the only eigenvalue of A, we have $\alpha + i\beta = 0$. Since H and K are Hermitian, the numbers α and β must be real. Hence $\alpha = \beta = 0$ and $\lambda - \alpha x - \beta y = \lambda$. But λ does not divide $f(x, y, \lambda)$. Thus, $f(x, y, \lambda)$ is an irreducible cubic. However, $(A^*)^2A^2$ has trace $\frac{1}{4}$, while $(B^*)^2B^2$ has trace 2, so A and B are not unitarily similar.

If A is a $3m \times 3m$ matrix and $f(x, y, \lambda) = [g(x, y, \lambda)]^m$, where $g(x, y, \lambda)$ is an irreducible cubic polynomial, then $m(x, y, \lambda) = g(x, y, \lambda)$, and Theorem 3 tells us that there is a unitary matrix U such that $U^*AU = A_1 \oplus \cdots \oplus A_m$, where each A_j is 3×3. If $A_j = H_j + iK_j$, then we must have $g(x, y, \lambda) = \det(\lambda I - xH_j - yK_j)$, but Example 2 shows that the A_js need not be unitarily similar.

If $m(x, y, \lambda)$ factors into linear factors, then H and K have property L, so $HK = KH$ and A is normal. If the cubic $m(x, y, \lambda)$ factors into a linear factor and an irreducible quadratic factor, and A is unitarily reduced into the finest possible block diagonal form, then Example 1 shows that there may or may not be 1×1 and 2×2 blocks. Theorem 2 of [9, p. 105] gives a sufficient condition for A to unitarily reduce to a sum of 1×1 and 2×2 blocks.

5. KIPPENHAHN'S CONJECTURE FOR $n \leqslant 5$.

Using the results of [8] and [9], and Theorem 3, we can show Kippenhahn's conjecture holds for $n \leqslant 5$. This was also done in [11] using a computational method. If $n \leqslant 5$ and $m(x, y, \lambda)$ has degree smaller than 5, then $f(x, y, \lambda)$ has either a repeated linear factor or a repeated quadratic factor. In the case of a repeated linear factor, the theorem of [8, p. 212] tells us H and K have a common eigenvector, and hence A is unitarily reducible. If there is a repeated quadratic factor, then $m(x, y, \lambda)$ has degree three and Theorem 3 guarantees A is unitarily reducible.

I would like to thank the referee for pointing out an error in the original proof of Theorem 3 and for offering helpful suggestions toward an alternative argument.

REFERENCES

1 G. P. Barker, L. Q. Eifler, and T. P. Kezlan, A non-commutative spectral theorem, *Linear Algebra Appl.* 20:95–100 (1978).

2 M. Fiedler, Geometry of the numerical range of matrices, *Linear Algebra Appl.* 37:81–96 (1981).

3 R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.* 6:193–228 (1951).

4 T. S. Motzkin and O. Taussky, Pairs of matrices with property L, *Trans. Amer. Math. Soc.* 73:108–114 (1952).

5 T. S. Motzkin and O. Taussky, Pairs of matrices with property L, II, *Trans. Amer. Math. Soc.* 80:387–401 (1955).

6 F. D. Murnaghan, On the field of values of a square matrix, *Proc. Nat. Acad. Sci. U.S.A.* 18:246–248 (1932).

7 F. Rellich, Störungstheorie der Spektralzerlegung, I, *Math. Ann.* 113:600–619 (1936).

8 H. Shapiro, A conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices, Part 1, *Linear Algebra Appl.* 43:201–221 (1982).

9 H. Shapiro, A conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices, Part 2, *Linear Algebra Appl.* 45:97–108 (1982).

10 H. Shapiro, Simultaneous block triangularization and block diagonalization of sets of matrices, *Linear Algebra Appl.* 25:129–137 (1979).

11 H. Shapiro, Unitary block diagonalization and the characteristic polynomial of a pencil generated by Hermitian matrices, Ph.D. thesis, California Institute of Technology, 1979.

12 W. Specht, Zur Theorie der Gruppen linearer Substitutionen, II, *Jahresber. Deutsch. Math.-Verein.* 49:207–215 (1939).

13 O. Taussky, A weak property L for pairs of matrices, *Math. Z.* 71:463–465 (1959).

Received 24 July 1981; revised 18 January 1982